PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/60, 1/21, 9/88
// (C12N 1/21, C12R 1/19)

A1

(11) International Publicati n Number: WO 90/10073

(43) International Publication Date: 7 September 1990 (07.09.90)

(21) International Application Number: PCT/CA90/00057

(22) International Filing Date: 21 February 1990 (21.02.90)

(30) Priority data:
314,879 24 February 1989 (24.02.89) US

(71) Applicant: MAJESTY (HER) IN RIGHT OF CANADA as represented by THE NATIONAL RESEARCH COUNCIL OF CANADA [CA/CA]; Montreal Road, Ottawa, Ontario K1A 0R6 (CA).

(72) Inventors: DE LUCA, Vincenzo ; 2513 Blain Avenue, Saskatoon, Saskatchewan S7J 2B7 (CA). BRISSON, Normand ; 6632 21e Avenue, Montreal, Quebec H1X 2G7 (CA). KURZ, Wolfgang, G., W. ; 514 Copland Crescent, Saskatoon, Saskatchewan S7H 2Z5 (CA).

(74) Agent: FETHERSTONHAUGH & CO.; 4, place Ville Marie, Suite 606, Montreal, Quebec H3B 2E7 (CA).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: A TRYPTAMINE PRODUCING TRYPTOPHAN DECARBOXYLASE GENE OF PLANT ORIGIN

(57) Abstract

Isolation and cloning of cDNA sequence of the tryptophan decarboxylase gene from *Catharanthus roseus* and the development of the cDNA sequence in a plasmid vector capable of transforming cell lines that will produce the tryptophan decarboxylase enzyme.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CM	Cameroon
DE	Germany, Federal Republic of
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
HU	Hungary
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
RO	Romania
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

TITLE OF THE INVENTION

A tryptamine producing tryptophan decarboxylase gene of plant origin.

BACKGROUND OF THE INVENTION

Tryptophan decarboxylase (TDC; E.C. 4.1.1.27) catalyses the conversion of L-tryptophan to tryptamine. This enzyme has been detected in numerous plant systems and it has been suggested that its primary role is to supply possible precursors for auxin biosynthesis (Baxter, C. & Slaytor, M. (1972) Phytochemistry 11, 2763-2766; Gibson, R.A., Barret, G. & Wightman F. (1972) J. Exp. Bot. 23, pages 775-786; Gross, W. & Klapchek, S. (1979) Z. Pflanzenphysiol. 93, pages 359-363).

In the Gramineae, TDC catalyses the synthesis of precursors for the protoalkaloids which have considerable physiological activity in higher animals (Smith, T.A., (1977) Phytochemistry Vol. 16, pages 171-175). It is also known that tryptophan-derived tryptamines are also precursors of the tricyclic β-carboline alkaloids formed by condensation with a one- or two- carbon moiety (Slaytor, M., & McFarlane, I.J., (1968) Phytochemistry 7, pages 605-610).

Furthermore, in periwinkle (Catharanthus roseus), TDC produces tryptamine for biosynthesis of the commercially important antineoplastic monoterpenoid indole alkaloids, vinblastine and vincristine (De Luca, V., & Kurz, W.G.W. (1988), Cell Culture and Somatic Cell Genetics of Plants, Constabel, F. and Vasil, I.K., eds. Academic Press 5, pages 385-401).

The TDC from Catharanthus roseus has been purified to homogeneity. It occurs as a dimer consisting of 2 identical subunits of Mr 54,000 and it requires pyridoxal phosphate for activity (Noe, W., Mollenschott, C., & Berlin J. (1984) Plant Mol. Biol. 3, pages 281-288).

The enzyme possesses characteristics of plant aromatic decarboxylases which usually exhibit high

substrate specificity. For example, TDC will decarboxylate L-tryptophan and 5-hydroxy-L-tryptophan but is inactive towards L-phenylalanine and L-tyrosine, while the tyrosine decarboxylases from Syringa vulgaris (Chapple, C.C.S., (1984) Ph.D. Thesis, University of Guelph, Guelph, Ontario, Canada), Thalictrum rugosum and Escholtzia californica (Marques, I.A., & Brodelius, P. (1988) Plant Physiol. 88, pages 52-55), accept L-tyrosine and L-dopa as substrates but not L-tryptophan or 5-hydroxy-L-tryptophan. The aromatic L-amino acid decarboxylases (dopa decarboxylase (DDC), ED 4.1.1.28) of D. melanogaster (Clark, W.C., Pass, P.S., Venkatararman, B., & Hodgetts, R.B. (1978) Mol. Gen. Genet. 162, pages 287-297; Eveleth, D.D., Gietz, R.D., Spencer C.A., Nargang, F.E., Hodgetts, R.B. & Marsh, J.L. (1986) Embo. J. 5, pages 2663-2672; Morgan B.A., Johnson, W.A. & Hirsh, J. (1986) Embo. J. 5, pages 3335-3342) and mammals (Albert, V.R., Allen, J.M., & Joh, T.H. (1987) J. Biol. Chem. 262, pages 9404-9411) have a broader substrate specificity with L-dopa, tyrosine, phenylalanine and possibly histidine also serving as substrates.

In animals, the role of aromatic L-amino acid decarboxylase is to produce the major neurotransmitters dopamine and serotonin and, in D. melanogaster, the DDC enzyme serves a second, inducible role, in the sclerotization of the insect cuticle (Christenson, J.G., Dairman, W. & Undenfriend, S. (1972) Proc. Natl. Acad. Sci. USA 69, pages 343-347; Lovenberg, W., Weissbach, W., & Undenfriend S. (1962) J. Biol. Chem. 237, pages 89-93; Yuwiler, A., Geller, E. & Eiduson, S. (1954) Arch. Biochem. Biophys. 80, pages 162-173; Brunet, P. (1980) Insect Biochem. 10, pages 467-500).

It would appear highly desirable to be able to clone the cDNA sequence of tryptophan decarboxylase from Catharanthus roseus, thus, providing the development of the cDNA sequence in a plasmid vector capable of

transforming cell lines that will produce the tryptophan decarboxylase enzyme.

If the tryptophan decarboxylase gene could be inserted into living organisms by transformation to produce tryptamine and related protoalkaloids, it could supplement a neurotransmitter deficiency.

Further, the insertion of this gene in plants could be useful to alter the spectrum of tryptophan-based chemicals normally produced by the plant. For example, the insertion of constitutive expression of tryptophan decarboxylase in Brassica species could sequester the cytoplasmic tryptophan pool for the synthesis of tryptamine and related protoalkaloids and therefore repress the normal synthesis and accumulation of indole glucosinolates.

Hence, creation of plants with an altered chemical spectrum may produce novel phenotypes which have resistance to various pathogenic diseases or to insect pests.

SUMMARY OF THE INVENTION

In accordance with the present invention, there is now provided the sequence of a cDNA clone which includes the complete coding region of tryptophan decarboxylase, preferably tryptophan decarboxylase (E.C. 4.1.1.27) from periwinkle (Catharanthus roseus). The cDNA clone (1747 bp) was isolated by antibody screening of a cDNA expression library produced from poly A^+ RNA found in developing seedlings of C. roseus. The clone hybridized to a 1.8 kb mRNA from developing seedlings and from young leaves of mature plants.

Also within the scope of the present invention is a method for inserting TDC gene into living organisms by transformation. The identity of the clone was confirmed when extracts of transformed E. coli expressed a protein containing tryptophan decarboxylase enzyme activity. The trypt phan decarboxylase cDNA clone encodes a protein of 500 amino acids with a calculated

molecular mass of 56,142 Da. The amino acid sequence shows a high degree of similarity with the aromatic L-amino acid decarboxylase (dopa-decarboxylase) and the alpha-methyldopa hypersensitive protein of Drosophila melagonaster. The tryptophan decarboxylase sequence also showed significant similarity to feline glutamic acid decarboxylase and mouse ornithine decarboxylase suggesting a possible evolutionary link between these amino acid decarboxylases.

Furthermore, the protein encoded by the cDNA clone of the present invention is active in vitro.

IN THE DRAWINGS

Figure 1 (lane 2) represents the TDC enzymatic activity in extracts of pTDC5-transformed E. coli, compared to those in control E. coli (lane 1) and that in C. roseus itself (lane 3).

Figure 2 represents the hybridization of the pTDC-5 clone to a 1.8 kb mRNA species isolated from periwinkle.

Figure 3 shows the nucleotide sequence of the pTDC5 cDNA clone and its deduced amino acid sequence. The putative polyadenylation signal is underlined.

Figure 4 shows the amino acid sequence alignments of the protein for the D. melanogaster alpha methyldopa hypersensitive gene (AMD), C. roseus tryptophan decarboxylase (TDC), and Drosphila DOPA decarboxylase isoenzyme 1 (DDC1).

Figure 5 shows hydropathy profile of TDC and DDC1.

Other advantages of the present invention will be readily illustrated by referring to the following description.

DETAILED DESCRIPTION OF THE INVENTION

cDNA synthesis and DNA sequencing.

Seedlings of C. roseus (L.) G. Don cv "Little Delicata" were germinated and grown for 5 days in the dark as described previously (De Luca, V., Alvarez-

Fernandez, F., Campbell, D., & Kurz, W.G.W. (1988) Plant Physiol. 86, 447-450). Seedlings were harvested after 18 hours of light treatment and total RNA was isolated as described by Jones, J.D.G., Dunsmuir, P. & Bedrook, J. (1985) EMBO J. 4, 2411-2418.

Poly(A)$^+$ RNA was isolated by chromatography on oligo (dT)$^-$ cellulose (Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412) and double-stranded cDNAs were prepared according to the procedure of Gubler and Hoffman (1983, Gene 25, 263-269). Following ligation with Eco RI linker, the cDNA was inserted into the Eco RI site of the expression vector ZAP (Stratagene, San Diego, Short, J.M., Fernandez, J.M., Sorge, J.A. & Huse, W.D. (1988) Nucl. Acids Res. 16, 7583-7600). A library containing 3.1 X 10^5 recombinant phages was obtained and after amplification, 2 X 10^5 plaques were screened with specific polyclonal antiserum raised against-TDC. Plasmids (pBluescript) containing a TDC cDNA insert were rescued using the R408 f1 helper phage (Short, J.M., Fernandez, J.M., Sorge, J.A. & Huse, W.D. (1988) Nucl. Acids Res. 16, 7583-7600) and the nucleotide sequence of a full-length cDNA clone (pTDC5) was determined on both strands by the dideoxy-chain-termination method (Sanger, F., Nicklen, S. & Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467). The sequencing strategy included subcloning of restriction fragments and the use of oligonucleotide primers. The sequence for all restriction sites used for the subcloning was determined on at least one strand. Comparisons of the pTDC5 cDNA nucleotide sequence and of the deduced amino acid sequence with Genbank and NBRF sequence libraries were performed using the FASTA program package (Pearson, W.R. & Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444-2448).

RNA blot hybridization.

Poly(A)$^+$ RNA was isolated from 6 day old developing seedlings and from y ung leaves of mature

plants as described above. These tissues were chosen as a likely source of TDC poly(A)$^+$ RNA based on the presence of high levels of TDC enzyme activity (De Luca, V., Alvarez-Fernandez, F., Campbell, D., & Kurz, W.G.W. (1988) Plant Physiol. 86, 447-450). RNA was denatured, fractioned by electrophoresis on formaldehyde/agarose gels, and then transferred to nitrocellulose filters (Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) In: Molecular Cloning, A Laboratory Manual. Cold Spring Harbor, New York). Blotted RNA was hybridized to [^{32}P]-labelled pTDC5 DNA and autoradiography was performed using Kodak XAR-5 films.

TDC activity in extracts of E. coli.

A culture (100 ml) of the E. coli strain ZL1-blue containing pTDC5 or pBluescript was incubated at 37^{0}C for 2 hours before adding the IPTG inducer at a final concentration of 1 mM. Incubation was continued for an additional 2 hours. Cells were harvested, washed in TE buffer, resuspended and lysed in 3 ml of a buffer containing 0.1 M Hepes, pH 7.5, 1 mM DTT. Debris was removed by centrifugation and the supernatant was desalted by passage over a Sephadex G-25^{TM} column. TDC enzymatic activity in bacterial supernatants was determined by monitoring the conversion of L-[methylene-^{14}C]-tryptophan to [^{14}C]-tryptamine (De Luca, V., Alvarez-Fernandez, F., Campbell, D., & Kurz, W.G.W. (1988) Plant Physiol. 86, 447-450). Supernatants (30 µl) were incubated in the presence of 0.1 µCi of [^{14}C]-tryptophan (sp. act. 59 mCi/mmol.) for 30 minutes and reactions were stopped with 100 µl NaOH. Radioactive tryptamine was extracted from the reaction mixture with ethyl acetate and was analyzed by silica gel thin layer chromatography and autoradiography. Determination of TDC enzyme activity in leaves was performed as described previously (De Luca, V., Alvarez-Fernandez, F., Campbell, D., & Kurz, W.G.W. (1988) Plant Physiol. 86, 447-450).

TDC enzymatic activity in E. coli.

A tryptophan decarboxylase cDNA clone of C. roseus was isolated by the use of antibody screening of an expression library. The antigenicity and enzymatic activity (Figure 1) of the encoded protein established the identity of the TDC cDNA.

When the original cDNA library was screened with the anti-TDC antibody, 27 clones were identified. Six clones were selected and submitted to further analysis. Partial sequence analysis revealed no difference among these clones, except for their length. Therefore, the clone having the longest cDNA insert (pTDC5) was selected for further characterization. To confirm that this cDNA clone corresponded to TDC, enzymatic activity was measured in cell extracts from E. coli. Figure 1 shows that [^{14}C]-tryptamine was produced with extracts from cells transformed with pTDC5, and with extracts from C. roseus leaves (lane 3), but not with extracts from cells containing only the vector (lane 1).

The conversion of [^{14}C]-tryptophan to [^{14}C]-tryptamine was monitored in extracts of E. coli and C. roseus leaves. [^{14}C]-tryptophan (sp. act. 50 mCi/mmol) for 30 minutes. After addition of base, [^{14}C]-tryptamine was extracted from the reaction mixture with ethyl acetate and reaction products were analyzed by thin layer chromatography on silica gel (solvent $CHCl_3$ MeOH: 25% NH_3 (5:4:1) and autoradiography. In Figure 1, TDC enzymatic activity is shown; lane 1, E. coli is transformed by the pBluescript vector, lane 2, E. coli is transformed by pTDC5 and lane 3, C. roseus extract is shown.

This result indicated that TDC enzymatic activity was retained by the protein produced using a TDC cDNA clone under the control of the Lac promoter of the pBluescript vector. No attempts were made to quantify the level of activity of TDC in E. coli.

Sequence analysis of a TDC cDNA clone.

DNA sequence analysis of pTDC5 revealed the presence of an open reading frame coding for a protein of 500 amino acids, which corresponded to a molecular mass of 56,142 Da (Figure 2). The 5'-nontranslated region of pTDC5 contained 69 nucleotides and included, near its beginning, a long stretch of alternating pyrimidines. Sequence around the methionine initiation codon (AAUAAUGGG) matched closely the consensus sequence for plant gene initiation codons (AACAAUGGC) (Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kerm, H.F. and Scheele, G.A. (1987) EMBO J. 6, 43-48). The 3'nontranslated region consisted of 168 nucleotides up to the poly(A) tail and contained an AAUAAA putative poly(A)$^+$ addition signal 17 nucleotides upstream from the start of the poly(A)$^+$ tail. Examination of the predicted amino acid sequence did not reveal the presence of a signal sequence (Watson, M.E.E. (1984) Nucl. Acids Res. 12, 5145-5164), which is consistent with the proposed cytoplasmic location of TDC within the cell (De Luca, V., Alvarez-Fernandez, F., Campbell, D., & Kurz, W.G.W. (1988) Plant Physiol. 86, 4474-50).

Comparison of TDC-cDNA nucleotide and deduced amino acid sequences with nucleotide sequences in the Genbank DNA sequence database and with amino acid sequences in the NBRF protein sequence database revealed surprising similarity (40% amino acid identity) with the dopa-decarboxylase isoenzyme 1(DDC1) from D. melanogaster (Eveleth, D.D., Gietz, R.D., Spencer, C.A., Nargang, F.E., Hodgetts, R.B. & Marsh, J.L. (1986) EMBO J. 5, 2663-2672; Morgan, B.A., Johnson, W.A. & Hirsh, J. (1986) EMBO J. 5, 3335-3342), and with the protein corresponding to the D. melanogaster alpha-methyldopa hypersensitive gene (AMD, 35% amino acid identity) (Eveleth, D.D. & Marsh, J.L. (1986) Genetics 114, 469-483) (Figure 3). In Figure 3, the boxes show TDC residues present in AMD and/or DDC1 sequences. Amino

acids are numbered for TDC (top) and DDC1 (bottom). The areas of amino acid similarity extended throughout the protein and were not restricted to a particular portion of either structure.

Furthermore, the 39% amino acid sequence similarity could be extended to the predicted distribution of potential alpha helices and beta sheets. This indicated that the amino acid differences between the two proteins did not significantly alter their secondary structures, and may indicate the importance of such conserved domains to mediate subunit assembly, as well as catalytic function and substrate specificity.

Limited proteolysis of pig kidney dopa decarboxylase and amino acid sequencing of a tryptic fragment produced a sequence for 50 amino acid residues one third of the distance from the COOH terminus of this protein (Tancini, B., Dominici, P., Simmaco, M., Schinina, M.E., Barra, D., & Voltatormi, C.D. (1988) Arch. Biochem. Biophys. 260, 569-576). Comparison of this 50 amino acid sequence with periwinkle TDC and D. melanogaster DDCI gave 20 and 32 identical amino acids, respectively. Furthermore, comparison of C. roseus TDC to feline glutamic acid decarboxylase (Kobayashi, Y., Kaufman, D.L. & Tobin, A.J. (1987) J. Neurosci. 7, 2768-2772) showed that 10% of the amino acid residues were identical between these two proteins. This similarity could be extended to 25% on a 396 aa stretch. Mouse ornithine decarboxylase (Kahana, C. & Nathans, D. (1985) Proc. Natl. Acad. Sci. USA 82, 1673-1677) showed a statistically significant (Pearson, W.R. & Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444-2448) 12% amino acid sequence similarity to the plant TDC which also extended throughout the protein sequence. We also found that the sequence Pro-His-Lys, beginning at position 317 in TDC, was identical to the sequence at the pyridoxal phosphate binding sites of D. melanogaster DDC (Marques, I.A., & Brodelius, P. (1988) Plant

Physiol. 88, 52-55; Clark, W.C., Pass, P.S., Venkataraman, B., & Hodgetts, R.B. (1978) Mol. Gen. Genet. 162, 287-297), feline glutamic acid decarboxylase (Kobayashi, Y., Kaufman, D.L. & Tobin, A.J. (1987) J. Neurosci. 7, 2768-2772) and pig dopa-decarboxylase (Bossa, F., Martini, F., Barra, D., Borri Voltatorni, C., Minelli, A. & Turano, C., (1977) Biochem. Biophys. Res. Commun. 78, 177-183). In contrast, the AMD protein, whose enzymatic function is unknown, contained the sequence Leu-His-Lys at the pyridoxal phosphate binding domain. The sequence similarity observed between TDC, feline glutamic acid decarboxylase and mouse ornithine decarboxylase also suggests an evolutionary link between these three amino acid decarboxylases.

Structural similarities between TDC and D. melanogaster DDC1 proteins were further revealed by comparing their hydropathy profiles (Figure 4). Each value was calculated as the average hydropathic index of a sequence of 9 amino acids and plotted to the middle residue of each sequence. Positive and negative values indicate hydrophobic and hydrophillic regions of the proteins, respectively. Close examination of the alignment of hydrophobic and hydrophillic regions for the two proteins showed a striking match between them, except for the area near the N terminus and the region around TDC residue 225.

Most decarboxylases require for their activity a pyridoxal phosphate co-factor linked to the C amino group of a lysine residue. The observed similarities around the pyridoxal binding site of pig kidney dopa-decarboxylase, D. melanogaster dopa-decarboxylase and feline glutamate decarboxylase with that of periwinkle TDC strongly suggests that lysine 319 of TDC binds pyridoxal phosphate.

The aromatic amino acid decarboxylases of plants, insects and mammals are remarkably similar in

subunit structure, molecular mass and kinetic properties (Maneckjee, R., & Baylin, S.B. (1983) Biochemistry 22, 6058-6063). Plant aromatic amino acid decarboxylases (Noe, W., Mollenschott, C. & Berline J. (1984) Plant Mol. Biol. 3, pages 281-288; Chapple, C.C.S., (1984) Ph.D. Thesis, University of Guelph, Guelph, Ontario, Canada; Marques, I.A., & Brodelius, P. (1988) Plant Physiol. 88, pages 52-55), in contrast to those from animals, display high substrate specificity for indole or phenolic substrates but not to both. The strong similarity observed between periwinkle TDC and DDC1 of D. melanogaster suggests that plant aromatic amino acid decarboxylase specific for tyrosine, phenylalanine or dihydroxyphenylalanine may be structurally similar to TDC and may, therefore, also be evolutionarily related. The recent purification of specific L-tyrosine decarboxylases (Marques, I.A., & Brodelius, P. (1988) Plant Physiol. 88, pages 52-55) to homogeneity should allow cloning of these genes and direct testing of this hypothesis.

TDC mRNA accumulation.

Total poly(A)$^+$ RNAs (1 µg) from six day old C. roseus seedlings and from young leaves of mature plants were run on an agarose/formaldehyde gel and were transferred to nitrocellulose paper. Hybridization was performed with [^{32}P]-labelled pTDC5 insert (sp. act. 1.2 X 10^8 cpm/µG). When total poly(A)$^+$ RNA isolated from six day old seedlings was probed with a 1.6 kb cDNA fragment isolated from pTDC5, a 1.8 kb mRNA was detected (Figure 5, lane 1). Young leaves from the mature plant also contained a 1.8 kb mRNA (Figure 5, lane 2). A fainter signal corresponding to a transcript of 3.2 kb was also present in both the lanes. This signal could be a precursor form of the TDC mRNA or an unrelated transcript having some sequence similarity to TDC.

CLAIMS

1. A DNA sequence comprising the cDNA sequence of the tryptophan decarboxylase gene.

2. The DNA sequence as defined in claim 1, wherein the tryptophan decarboxylase gene is closed and sequenced from Catharanthus roseus.

3. A synthetic recombinant DNA molecule containing a DNA sequence comprising the cDNA sequence of the tryptophan decarboxylase gene.

4. A synthetic recombinant DNA molecule as defined in claim 3, wherein the tryptophan decarboxylase gene is cloned and sequenced from Catharanthus roseus.

5. A synthetic DNA molecule expressible in E. coli and coding for the expression of the tryptophan decarboxylase enzyme.

6. The synthetic DNA molecule of claim 5, wherein the tryptophan decarboxylase gene is from Catharanthus roseus.

7. An expression vector comprising a synthetic DNA molecule coding for the tryptophan decarboxylase enzyme.

8. An expression vector having a microorganism replication system and a gene coding for the expression of the tryptophan decarboxylase enzyme.

9. The expression vector of claim 8, wherein the microorganism is E. coli and wherein the tryptophan decarboxylase enzyme is from Catharanthus roseus.

10. A host cell having an extrachromosomal functional synthetic gene expressing an active tryptophan decarboxylase enzyme.

11. A cell according to claim 10, wherein said cell is a microorganism and wherein said tryptophan decarboxylase enzyme is from Catharanthus roseus.

12. A cell according to claim 11, wherein said microorganism is a bacterium.

13. A cell according to claim 11, wherein said bacterium is E. coli.

14. An E. coli bacteria having an extrachromosomal functional synthetic gene expressing an active tryptophan decarboxylase enzyme.

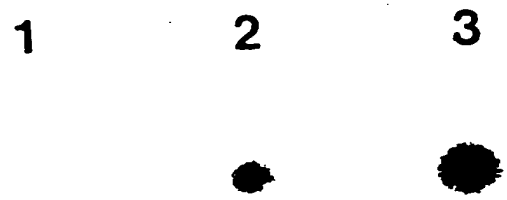

FIG - 1

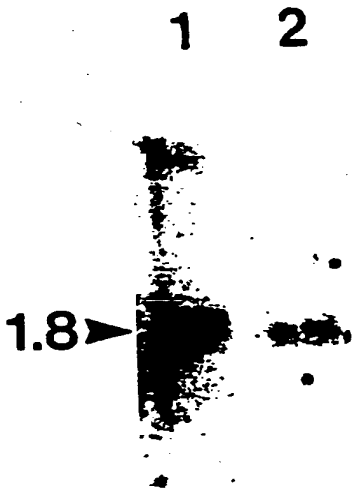

FIG - 2

```
1     CTCTCTCTCTCTCTCTCTAAGACTTTCTCTCTCTACACATACACCTAC
      asn val ala met ser asn ser pro val gly glu phe
91    AAT GTA GCC ATG TCC AAT TCT CCA GTT GGA GAA TTT
      arg met val asp phe ile ala asp tyr tyr lys asn
166   CGT ATG GTA GAT TTC ATA GCC GAT TAT TAC AAA AAT
      tyr leu arg lys arg ile pro glu thr ala pro tyr
241   TAT CTC CGA AAA CGT ATC CCC GAA ACC GCT CCT TAC
      lys asp ile ile pro gly met thr asn trp met ser
316   AAG GAT ATT ATC CCA GGA ATG ACA AAT TGG ATG AGC
      ala ala phe leu gly glu met leu ser thr ala leu
391   GCT GCC TTT TTA GGA GAA ATG TTG TCT ACT GCC CTA
      thr glu leu glu met ile val met asp trp leu ala
466   ACC GAA TTA GAA ATG ATT GTT ATG GAT TGG TTG GCT
      thr gly gly gly val ile gln asn thr thr ser glu
541   ACC GGT GGC GGC GTC ATC CAA AAC ACC ACT AGC GAG
      leu glu lys leu gly pro asp ser ile gly lys leu
616   CTG GAG AAG CTC GGT CCC GAT AGT ATT GGA AAA CTT
      lys thr cys lys leu ala gly ile tyr pro asn asn
691   AAA ACT TGC AAA TTG GCG GGA ATT TAT CCG AAT AAT
      ile ser pro gln val leu arg lys met val glu asp
766   ATC TCA CCT CAA GTT CTA CGA AAA ATG GTC GAG GAT
      thr leu gly thr thr ser thr thr ala thr asp pro
841   ACC CTG GGT ACC ACC TCG ACC ACG GCT ACC GAT CCT
      trp ile his val asp ala ala tyr ala gly ser ala
916   TGG ATC CAC GTG GAT GCT GCT TAT GCG GGA AGC GCC
      glu arg val asp ser leu ser leu ser pro his lys
991   GAA CGA GTT GAC TCA CTG AGT CTG AGT CCA CAC AAA
      lys gln pro his leu leu leu arg ala leu thr thr
1066  AAG CAA CCA CAT TTG TTA CTA AGG GCA CTC ACT ACG
      lys val val asp phe lys asn trp gln ile ala thr
1141  AAA GTT GTG GAC TTC AAA AAT TGG CAA ATC GCA ACG
      arg ser tyr gly val val asn leu gln ser his ile
1216  CGT AGC TAT GGA GTT GTT AAT TTA CAG AGT CAT ATT
      val arg ser asp ser arg phe glu ile val val pro
1291  GTT AGA TCA GAC TCC AGA TTC GAA ATT GTG GTA CCG
      val ser ser leu his val glu glu val asn lys lys
1366  GTT TCG AGT TTA CAT GTA GAA GAA GTG AAT AAG AAA
      thr his thr ile val gly gly ile tyr met leu arg
1441  ACT CAT ACT ATT GTG GGA GGC ATA TAC ATG CTA AGA
      arg arg val trp asp leu ile gln lys leu thr asp
1516  CGC CGT GTT TGG GAT TTG ATT CAA AAA TTA ACC GAT
1597  ATTTTTTTTTAAATTTTATATTTGCTGATTGTTTGAAGAGTTTAAAAA
1697  ATGTATTAATTATGACATGAGAATAAAATAGAATTTGTGTGTGCAAA
```

Fig. 3

```
                                        met  gly  ser  ile  asp  ser  thr
ACCAGAAAAAAGAAAAAAATA                   ATG  GGC  AGC  ATT  GAT  TCA  ACA
 lys  pro  leu  glu  ala  glu  glu  phe  arg  lys  gln  ala  his
 AAG  CCA  CTT  GAA  GCT  GAG  GAA  TTC  CGA  AAA  CAA  GCC  CAT
 val  glu  thr  tyr  pro  val  leu  ser  glu  val  glu  pro  gly
 GTG  GAA  ACA  TAT  CCG  GTC  CTT  AGC  GAA  GTC  GAA  CCT  GGA
 leu  pro  glu  pro  leu  asp  asp  ile  met  lys  asp  ile  gln
 CTC  CCC  GAA  CCA  CTT  GAC  GAC  ATC  ATG  AAA  GAT  ATT  CAG
 pro  asn  phe  tyr  ala  phe  phe  pro  ala  thr  val  ser  ser
 CCT  AAT  TTT  TAT  GCA  TTT  TTT  CCT  GCC  ACT  GTT  AGT  TCA
 asn  ser  val  gly  phe  thr  trp  val  ser  ser  pro  ala  ala
 AAT  TCA  GTA  GGC  TTT  ACT  TGG  GTT  TCT  TCA  CCA  GCC  GCC
 gln  ile  leu  lys  leu  pro  lys  ser  phe  met  phe  ser  gly
 CAG  ATC  CTT  AAA  CTC  CCC  AAA  TCT  TTC  ATG  TTT  TCA  GGT
 ser  ile  leu  cys  thr  ile  ile  ala  ala  arg  glu  arg  ala
 TCC  ATT  CTT  TGT  ACA  ATC  ATT  GCC  GCC  CGG  GAA  AGG  GCC
 val  cys  tyr  gly  ser  asp  gln  thr  his  thr  met  phe  pro
 GTC  TGT  TAC  GGA  TCC  GAT  CAA  ACC  CAT  ACC  ATG  TTC  CCC
 ile  arg  leu  ile  pro  thr  thr  val  glu  thr  asp  phe  gly
 ATT  AGG  TTA  ATA  CCT  ACG  ACC  GTC  GAA  ACG  GAT  TTC  GGC
 asp  val  ala  ala  gly  tyr  val  pro  leu  phe  leu  cys  ala
 GAC  GTG  GCG  GCC  GGA  TAT  GTA  CCG  CTG  TTC  TTA  TGC  GCT
 val  asp  ser  leu  ser  glu  ile  ala  asn  glu  phe  gly  ile
 GTG  GAC  TCA  CTT  TCT  GAA  ATC  GCT  AAC  GAG  TTT  GGT  ATT
 cys  ile  cys  pro  glu  phe  arg  his  tyr  leu  asp  gly  ile
 TGT  ATA  TGT  CCC  GAG  TTT  AGA  CAT  TAC  TTG  GAT  GGA  ATC
 trp  leu  leu  ala  tyr  leu  asp  cys  thr  cys  leu  trp  val
 TGG  CTA  CTC  GCT  TAC  TTA  GAT  TGC  ACT  TGC  TTG  TGG  GTC
 asn  pro  glu  tyr  leu  lys  asn  lys  gln  ser  asp  leu  asp
 AAT  CCT  GAG  TAT  TTA  AAA  AAT  AAA  CAG  AGT  GAT  TTA  GAC
 gly  arg  lys  phe  arg  ser  leu  lys  leu  trp  leu  ile  leu
 GGA  CGA  AAA  TTT  CGG  TCG  CTG  AAA  CTT  TGG  CTC  ATT  TTA
 arg  ser  asp  val  ala  met  gly  lys  met  phe  glu  glu  trp
 CGT  TCT  GAC  GTC  GCA  ATG  GGC  AAA  ATG  TTC  GAA  GAA  TGG
 arg  asn  phe  ser  leu  val  cys  phe  arg  leu  lys  pro  asp
 AGA  AAC  TTT  TCT  CTT  GTT  TGT  TTT  AGA  TTA  AAA  CCT  GAC
 leu  leu  asp  met  leu  asn  ser  thr  gly  arg  val  tyr  met
 CTT  TTG  GAC  ATG  CTT  AAC  TCG  ACG  GGA  CGA  GTT  TAT  ATG
 leu  ala  val  gly  ser  ser  leu  thr  glu  glu  his  his  val
 CTG  GCT  GTT  GGC  TCA  TCG  CTA  ACT  GAA  GAA  CAT  CAT  GTA
 asp  leu  leu  lys  glu  ala  ter
 GAT  TTG  CTC  AAA  GAA  GCT  TGA  TGAATAAGTAAGGGTTTTTTTTTA
TAAAGTGATTTGTAAAGGTTTATTGTACTCAAACAATCATGCAATTAATTAT
AAAA
```

FIG. 3

(CONT.)

```
                10        20        30        40        50        60
AMD  MGSSHDSTNVAMSNSPVGEFFPL----------EAEEFRKQAHRMVDFIADYLENIR
TDC  MSHHPISNTIPTKQTDGNGKANISPDKLDPKVSIDMEAPEFRDFAKTMVDFINEYLENIR
DDC1           10        20        30        40        50        60

                60        70        80        90       100       110
AMD  DDDVLPNVEPGYLDLLPTEMPEEPEAWKDVLGDISRVIMPGLTHSESPHMIAYYPTSTS
TDC  TYPVLSEVEPGYLRKRIPETAPYLPEPLDDIMKDIQDIIMPGMTNWMSPNFMAFFPATVS
DDC1 ERRVLPEVRPGYLKPLIPDAAPERPEEWQDVMNDIERVIMPGVTHWHSPKFHAYYPTANS
                70        80        90       100       110       120

                120       130       140       150       160
AMD  YPSIVGEMLASGFGVIGFSWICSPACTELEVVVMDWLAKFLKLPAHFLNGHASDGGGVIQ
TDC  SAAPLGEMLSMALNSVGFTWVSSPAATELEMIVMDWLAQILKLPKSFM--FSGTGGGVIQ
DDC1 YPAYVADMLSGAIACIGFTWIASPACTELEVVVMDWLGKMLELPAEFLACSGKGGGVIQ
                130       140       150       160       170       180

                180       190       200       210       220
AMD  GSASEAVLVAVLAARDQALANYRESHPEL-SESEVRGRLVAYSSDQSNSCIEKAGVLAAM
TDC  NTTSESILCTIIAARERALE----------KLGPDSIGKLVCYGSDQTHTMFPKTCKLAGI
DDC1 GTASESILVASAGSQGQEVE-GEGAPSGVGLEHTILGKLVGYCSDQAHSSVERAGLLGGV
                190       200       210       220       230       240

                240       250       260       270       280       290
AMD  ---PIRLLPA--GEDFVLRGDTLRGAIEEDVAAGRIPVICVATLGTTGTTDAMDDIESLSA
TDC  YPNNIRLIPTTVETDFGISPQVLRRMVEDVAAGVVPLFLCATLGTTSTTAIDPVDSLSE
DDC1 ---KLRSVQSENHR---MRGAALEKAIEDVAEGLIPFYAVVTLGTTNSCAFDYLDECGP
                250       260       270       280       290
```

FIG. 4

```
                280       290       300       310       320       330
AMD  VCEEFKCGSMLMPRMRWSLCSGGMF-GFAKGIGSRGLARLQPAQVHAGGLRLLGHVARGC
TDC  IANEEGIWIHVDAAYAGSACICPEFRHYLDGIERVDSLSLSPHKWLLAYLDCTCLWVKQP
DDC1 VGNKHNLWIHVDAAYAGSAFICPEMRHLMKGIESADSFNFNPHKWMLVNFDCSAMWLKDP
                300       310       320       330       340       350

                340       350       360       370       380       390
AMD  QQGGRQLQCGSHLSEAQARGSVANSRLPSLANPLGRRFRALRVMITFRTLEAEGLRNHVA
TDC  HLLIRALITNPEYLRNKQSDLDKVVDFKNWQIATGRKFRSLKLWLILRSYGVVNLQSHIR
DDC1 SWVVNAFNVDPLYLRHDMQ--GSAPDYRHWQIPLGRRFRALKLWFVLRLYGVENLQAHIR
                360       370       380       390       400       410

                400       410       420       430       440       450
AMD  KHIELAKQFEQLVLKDSRFELVAPRALGLVCFRPKGD-------NEITTQLLQRLMDTKRV
TDC  SDVAMGKMFEEWVRSDSRFEIVVPRNFSLVCFRLKPDVSSLHVEEVNKKLLDMLNSTGRV
DDC1 RHCNFAKQFGDLCVADSRPELAAEINMGLVSFRLKGS-------NERNEALLKRLINGRGHI
                420       430       440       450       460

                460       470       480       490       500
AMD  YMTHKTEHAGRQFLRFVVCGMDTKASDIDFAMQEIESQLTDLLQADESLVARKSGNVGDLAH
TDC  YMTHQTIVGGIYMLRLAVGSSLTEEHHVRRVWDLIQKLTDDLLKEA
DDC1 HLVPAKIKDVYGLRMAICSRFTQSEDMEYSWKEVSAAADEMEQEQ
                470       480       490       500       510

AMD  DFQIHLSTENATHEKSQ
```

FIG. 4
(CONT.)

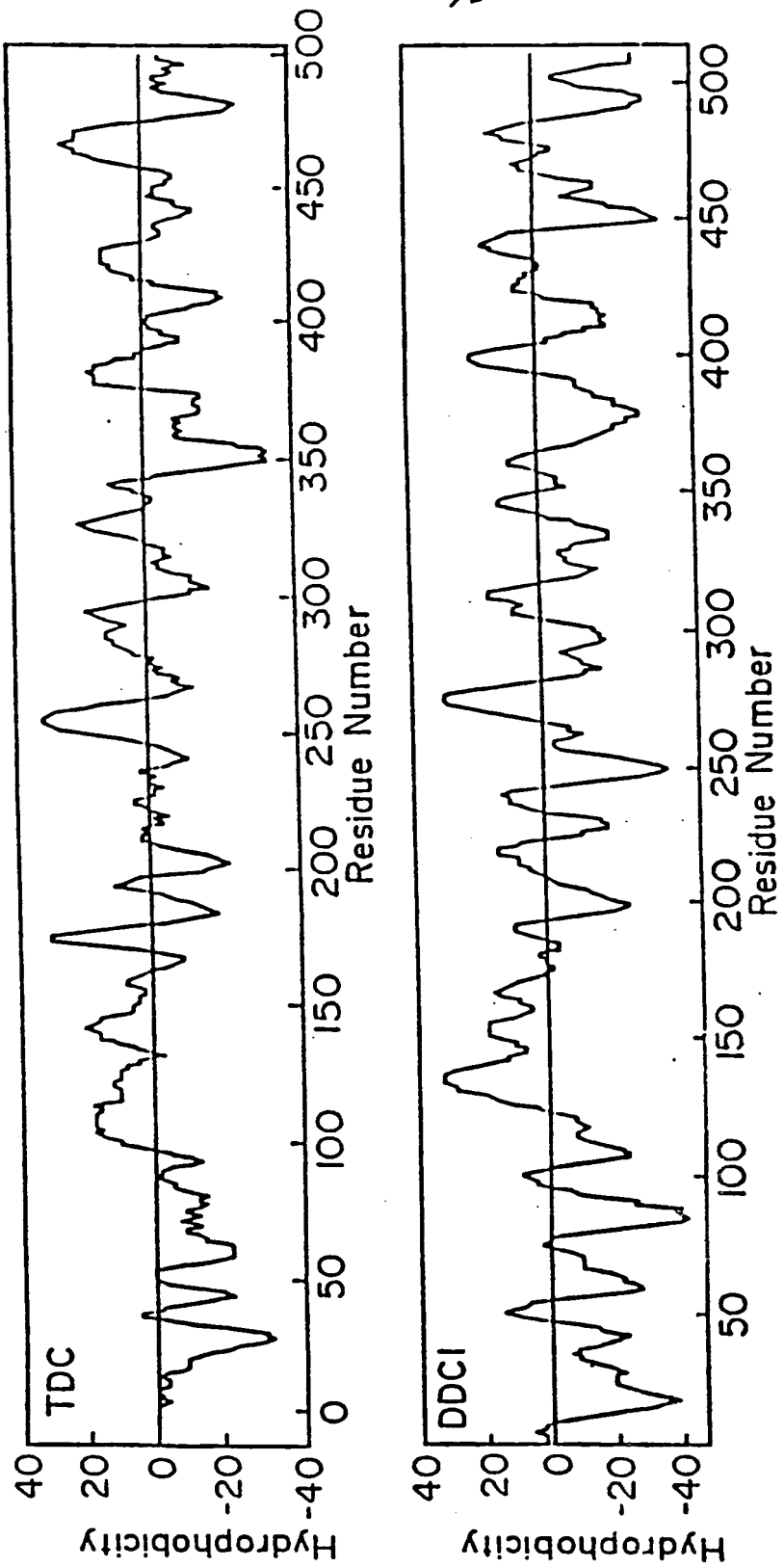
TDC
DDCI
Hydrophobicity
Residue Number
40
20
0
-20
-40
0
50
100
150
200
250
300
350
400
450
500

FIG. 5

INTERNATIONAL SEARCH REPORT

International Application No PCT/CA 90/00057

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/60, 1/21, 9/88, // (C 12 N 1/21, C 12 R 1:19)

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y	Plant Molecular Biology, no. 3, 1984, Martinus Nijhoff/Dr W. Junk Publishers, (Dordrecht, NL), W. Noé et al.: "Tryptophan decarboxylase from Catharanthus roseus cell suspension cultures: purification, molecular and kinetic data of the homogenous protein, pages 281-288 see the whole document cited in the application --	1-14
Y	EMBO Journal, volume 5, no. 10, 1986, IRL Press Limited, (Oxford, GB), D.D. Eveleth et al.: "Sequence and structure of the dopa decarboxylase gene of Drosophila: evidence for novel RNA splicing variants", see pages 2664-2668 cited in the application --	1-14

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
16th May 1990	15. 06. 90
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	M. Peis / M. PEIS

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	Progress in Catecholamine Research, Part A: Basic Aspects and Peripheral Mechanisms, 1988, Alan R. Liss, Inc., Tong H. Joh et al.: "Molecular biology of Aromatic L-amino acid decarboxylase and dopamine beta-hydroxylase", see pages 30-31 --	1-14
A	Journal of Biological Chemistry, volume 262, no. 19, 5 July 1987, The American Society of Biological Chemists, Inc., (US), V.R. Albert et al.: "A single gene codes for aromatic L-amino acid decarboxylase in both neuronal and non-neuronal tissues", pages 8404-9411 see page 9406 --	1-14
P,X	Procl. Natl. Acad. Sci. USA, volume 86, April 1989, Biochemistry, (US), V. De Luca et al.: "Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: Comparison with animal dopa decarboxylases", pages 2582-2586 see the whole document --------	1-14